Evaluation of the Impact of the Provision of Triple Combination Antiretroviral Therapy to Employees Through the In-House Health Programme at a Large South African Mining Company

Jonathan Edward Brink
BRNJON019

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN
In partial fulfilment of the requirements for the degree

Master of Commerce specialising in Economics

Faculty of Commerce
UNIVERSITY OF CAPE TOWN

Date of submission: 10 February 2018
Supervisor: Martin Wittenberg

School of Economics

With Thanks and Appreciation to

Anglo American Coal South Africa,

Dr Jan Pienaar,
Dr Brian Brink,
Ms Thia Grobler,
Dr Declan French

And

Prof Martin Wittenberg

INTRODUCTION

HIV/AIDS poses a unique challenge to businesses, particularly those operating in Southern Africa. The region is home to one third of the worldwide HIV positive population (as measured by those aged 15 to 49) yet as a whole contributes a diminutive proportion of the total worldwide population by the same measure (UNAIDS, 2016; World Bank, 2016).
Relative to its size, the epidemic has introduced disparately large economic strain due to the fact that the highest HIV prevalence rates coincide with the most productive years of people's lives, with prevalence rates peaking around the 30 to 34-year-old stratum (Shisana, et al., 2012). Loss of business productivity as a result of HIV related illness through a combination of absenteeism and so called 'presenteeism'[1] as well as death due to AIDS and the resultant increase in employee turnover, has motivated companies to implement workplace HIV education, prevention and treatment programmes over and above governments' efforts to curb the effect of the disease on business operations and the associated economic costs (Granich et al., 2012; Meyer-Rath et al., 2012, 2015).
As Rosen et. al. argue, HIV/AIDS is impacting two of the fundamental pillars behind the rationale driving investment in developing countries – that is; cheap labour and rapid market growth (Rosen, et al., 2003).

Fortunately, since the advent of effective antiretroviral therapy (ART), coupled with the massive drop in treatment costs over the years, HIV is no longer a death sentence, nor does the virus, under suppression, cause noticeable direct long-term disability if managed effectively (Deeks, Lewin, & Havlir, 2013; Menzies et al., 2011). This has introduced the opportunity for businesses to intervene with comprehensive treatment and care programmmes for their HIV positive employees in order to fight back at an epidemic that would otherwise be crippling.

At the time of any given employee contracting HIV, and given a no-intervention-scenario, a company can immediately conceptualize a contingent liability in the form of loss in productivity, increase in medical costs, increase in absentee days and death benefit payments should the employee die from AIDS. In order to mitigate these risks, companies ought to provide access to ART, following which, the contingent liability of AIDS related events should fall dramatically when compared to a no-treatment-scenario (Freedberg et al., 2001; INSIGHT START Study Group, 2015). HIV has become an unavoidable concern for businesses operating in high HIV prevalence countries, and so, the provision of ART must be assessed as a prudent investment decision rather than merely a stream of costs (Rosen, et al., 2004), and as with any investment should be acted on based on its merits. This paper aims to quantify the benefits of the provision of ART to HIV positive employees, as well as employers, through the analysis of the in-house HIV management programme implemented at Anglo American Coal South Africa (AACSA).

The data used to conduct this analysis come from the proprietary health information system ('theHealthSource') developed at AACSA in an effort to better manage health outcomes, treatment efficacies, as well as allow for the provision of information necessary for health programme evaluations. AACSA has provided access to its comprehensive data base of its workforce's health, anonymously linking HIV treatment statuses, and HR records, dating back to 2009. The exclusivity of these type of data, in any such scenario, is what makes this analysis unique among most. The Anglo American group of companies is at the forefront of company-level HIV treatment and monitoring, thus providing few scenarios in which the base of this analysis could be improved upon.

The monitoring and evaluation of these types of programmes plays an instrumental role in the advocacy of in-house HIV treatment strategies, and perhaps treatment of chronic illness as a whole.
What follows in this paper is an analysis of the company-level ART programme with the aim of gaining an understanding into a number of key factors through which ART and company level treatment programmes have their largest impact.

ETHICAL CONSIDERATIONS

The data used in this analysis were taken from a system that has generated unique random identifiers per employee, with no way to trace these identifiers back to any individual within the company. For the majority of the data, the patients had explicitly agreed to make available their disidentified data for monitoring and evaluation purposes. The individual level data within the company have been kept confidential at every level beyond the direct and specific use of it for medical treatment and care. Ethical approval for this study was granted by the University of Cape Town's Ethics committee (2016).

LITERATURE REVIEW

Background

Worldwide, South Africa, with an HIV positive population of 7,1 million people, is the country most heavily burdened by the epidemic (UNAIDS, 2016). As of 2016 18,9% of the overall working-age population is living with the disease. In many sub populations, however, the HIV prevalence rate reaches much higher levels (Shisana et al., 2014). The high prevalence rate, coupled with the fact that, historically, the background incidence of infectious diseases is generally high in developing countries such as South Africa, creates for an incubator in which opportunistic infections are able to thrive (Corbett et al., 2003).

Beyond the cost of HIV and AIDS to individuals, companies share in the burden of the disease in the form of increased costs attributable to heavy productivity losses, increased staff turnover due to ill health retirements and AIDS related mortality, as well as increased general healthcare programme costs. For this reason, companies, over and above the public sector, share in the incentive to provide antiretroviral therapy in order to combat the effects of the epidemic.

Antiretroviral therapy has been shown to not only drastically prolong the lives of HIV positive individuals but also

[1] Presenteeism refers to a reduction in on-the-job productivity usually attributed to illness (Hemp, 2004).

significantly reduce the risk of transmission of the virus from those on treatment with suppressed viral loads to HIV negative people (Deeks, Lewin, & Havlir, 2013; Menzies et al., 2011; Attia, Egger, Müller, Zwahlen, & Low, 2009).

The hidden threat of HIV for businesses lies in the somewhat unique pathogenesis of the virus in which HIV positive people can live asymptomatic and otherwise completely normal and healthy lives for 5 to 10 years following initial infection (Rosen, et al., 2004)(Kenya). The virus, thus, has the ability to penetrate populations with few, if any, short term consequences yet cause deep structural damage to the long-term prospects of any business or economy in which it infiltrates (Poku, 2004). For this reason it can be hard to convince management of the necessities of tackling a problem which they cannot see, or at the very least, and for all purposes, is considered to be a distant problem. As Rosen et. al. argue, the cost of HIV/AIDS for employers must be based on prevalence and not incidence – that is, the discounted cost of an HIV infection must be measured at the time of contraction of the virus and not at the time symptoms appear if organisations are to justify the cost of prevention and treatment (Rosen, Simon, Thea, & Vincent, 2000).

HIV has vastly differing prevalence rates and distributions across regions, genders, age, ethnic groups, sexual orientations as well various other sub-populations such as migrant labour (Shisana et al., 2014) (Brummer, 2002).By virtue of this, the epidemic also has vastly differing implications for businesses dependent on these sub-populations (Barnett & Whiteside, 1999). Historically, migrant labour has been a particularly significant risk factor in the susceptibility to HIV infection (Lurie, 2006), driven, in part, by the fact that these workers live away from their families, often in single sex hostels, with affordable access to alcohol and sex workers – all of which contribute to risky sexual behaviours (Crush, Williams, Gouws, & Lurie, 2005).

HIV EPIDEMIOLOGY

If HIV is allowed to progress to full-blown AIDS, there is a dramatic increase in the incidence, length and severity of opportunistic infections, chief of which and overall the deadliest for those suffering with AIDS, is tuberculosis (Barnett & Whiteside, 2000)(Southern Africa) – a disease which has accelerated off the back of the HIV epidemic. Among most of the Southern African countries the proportion of people with incident TB infections who are HIV positive sits at more than 60% (World Health Organisation, 2016), with HIV coinfection of incident TB cases in South Africa sitting at above 70% (SANAC, 2011). This is of particular concern for mining companies whose employees, beyond just being at a higher risk for HIV infection through the exposure to a number of risk factors (Lurie, 2006), are, under some circumstances, exposed to inhaled particulates such as crystalline silica dust which can create a predisposition to TB infection through the development of pneumoconiosis (Goldstein & Webster, 1972; Kim et al., 2009). The triad of HIV, TB and pneumoconiosis creates a particularly devastating burden of disease in the mining industry.

In Sub-Saharan African countries, where HIV-TB coinfection rates are abnormally high, the provision of ART has been shown to be a cost-effective means of TB management (Williams, et al., 2010), and, indeed, a highly effective means of management of all opportunistic infections. Anecdotal evidence shared by Anglo American Coal South Africa suggests that their response in rolling out HIV treatment has drastically reduced the annual incidence of opportunistic TB infections, although the causality of this relationship is yet to be established, and generally beyond the scope of our analysis, in which we look primarily at the relationship between HIV treatment and company reported sick days, as well as clinical visits per employee.

The primary mechanism of clinical action of ART is in the reduction of HIV viral loads and the resultant preservation of healthy CD4 counts in HIV positive individuals. Generally, if treatment is working, those on ART will have much higher CD4 counts as well as lower viral loads than those not on ART when conditioned on the length of time from HIV infection (Battegay, Nüesch, Hirschel, & Kaufmann, 2006), and when used effectively for those presenting with varying degrees of immunosuppression, has the potential to allow recovery of CD4 counts in treated HIV positive people to the point where CD4 counts are normalized and comparable to those observed in HIV negative groups (Mocroft et al., 2007). Those not suffering the effects of immunosupression under similar circumstances to those with varying degrees of HIV induced immunosuppression are far less likely to contract opportunistic infections (Holmes et al., 2006; Lloyd, 1996).

PRIOR RESEARCH ON THE IMPACT OF HIV AND ART ON LABOUR OUTCOMES

A central theme in the literature that has been built around the analysis of the impact of HIV and related treatment of the disease in individuals is the decline in health states and/or the related loss in productivity building up to dates of separation from employment. These separations could be as a result of either ill-health early retirements or death. Other possible analyses on the effects of treatment on individuals are based on the date of first treatment through the use of ART. Whether related to employment outcomes, health decline or the effect of treatment, the analyses are focused mainly on the patterns of observations of various outcome variables around a critical event date, which in keeping with the literature, we will refer to as date 0. A large motivation for the use of an event date as a reference point in which estimations are based around (other than that of the date of infection) in the analysis of outcomes related to HIV is the vast heterogeneity in the timing of disease progression and related asymptomatic period of infection. By looking at the patterns around an event date, such as separation from employment or date of first treatment, one can implicitly control for the varied rates of disease progression across patients. The patterns for each individual would be similar around death, ill-health retirement, or treatment dates regardless of how long it takes for each individual to get to the symptomatic stage of the disease in which a death, an ill-health retirement, or treatment occurs.

In a seminal paper by Fox et al. (2004), (Kenya) in which tea estate employees' productivity could be directly observed through the measurement of the weight of tea leaves plucked per day per individual, it was found that there was an increase in the number of leave days for an HIV positive worker in the three years prior to an AIDS related ill-health retirement or death.

HIV positive employees in the year prior to termination took on average around 10 more sick leave days than the control group with a similar difference in the number of casual and annual leave days, too. Productivity also declined by an average of 17,7% in the year prior to termination.

In a study later conducted by Larson et al. in 2013 (Kenya), the authors analysed not only the decline in health states and productivity with HIV disease progression but also the reversal of this trend following ART initiation. The study focused on the same tea estate as that which was reported on in Fox et al (2004). The authors used nearest neighbour matching techniques in order to match individuals to controls in the HIV negative population. It was found that in the month of ART initiation, male and female cases plucked 51% and 62% less tea leaves respectively than the reference group and spent on average 47% and 57 % fewer days plucking leaves. After 2 years on ART, male and female cases were on average 8% and 19% less productive than the matched controls.

Sonnenberg et al. (2011) analysed a large cohort of South African gold miners, finding that HIV positive workers in the year before death had a 38,8% absenteeism rate. The authors also approximate the increased demand for medical services by analysing absences due to medical reasons to which a 13,6% absenteeism rate could be ascribed.

In an early analysis of Anglo American's HIV treatment Programme, Muirhead et al. (2006) looked into the viability of public private partnerships in increasing access to ART. The authors noted that HIV positive employees took on average 7,5 sick leave days in the month prior to ART initiation, which fell to 2,1 days after 18 months of treatment.

Habyarimana, Mbakile, & Pop-Eleches (2010) use data spanning from 2001 to 2006 at the Debswana goldmine in Botswana to estimate patterns of per person absenteeism around ART initiation dates. They find that HIV positive employees take roughly an extra 5 sick days in the month of treatment initiation when compared to exactly 12 months prior to treatment initiation. Further to this, the authors plot out a modelled time path of disease progression in the absence of treatment for each treated individual. This estimated time path allows for a 'true' counterfactual to the treated case to be calculated for each individual.
The discussion and estimations henceforth draw influences from the above research, most notably those papers by Habyarimana, Rosen and Larson in an attempt to estimate patterns of labour and health outcomes around treatment initiation dates at Anglo American Coal South Africa.

ANGLO AMERICAN HEALTH PROGRAMME DESIGN AND MANAGEMENT

As one of the first extensive workplace HIV treatment programmes in the world, the Anglo American group of companies' HIV treatment programme design has been well documented in the paper 'Evaluation of a Workplace HIV Treatment Programme in South Africa' (Charalambousa, et al., 2007). There have been few fundamental changes to the programme design over the years other than the criteria stipulating eligibly for HIV treatment, which has advanced from a model of that based on CD4 counts and/or diagnosis of severe opportunistic infections to a universal test and treat model[2]. With reference to the prior work by Charalambousa et al. (2007), and in further discussion with managers at the company, a detailed description of the history and policies at the company follows:

A distinction is drawn between the enrolment in the HIV programme and HIV treatment uptake. Those diagnosed with HIV through a voluntary counselling and testing session are offered specialized HIV care through the company. Those who opt not to seek care through the company can either seek care elsewhere or refuse care entirely. HIV programme enrolment offers quarterly health checks which include screenings for opportunistic infections, CD4 count checks and HIV viral load checks in order to monitor disease progression. Although, across the entire sample period, any given patient may have become eligible for treatment, treatment uptake has always been a voluntary decision for the patient.

Originally the company's policy, at a patient level, on the in-house provision of ART was based on eligibility for treatment by CD4 count, or diagnosis of severe opportunistic infections – those with CD4 counts below 350 cells per millilitre of blood were initiated on treatment without further delay. The company has since moved to a universal test-and-treat model in which anyone diagnosed with HIV becomes eligible for treatment as of the diagnosis date.
There are several motivations driving this strategy.

1) The proportion of the total additional cost over a patient's lifetime of a universal test-and-treat strategy over one based on CD4 counts is minimal.
2) Waiting for CD4 counts to drop below a threshold level puts that person at risk of
   a) waiting too long to test CD4 counts again
   b) increasing the risk of contracting an opportunistic infection, such as TB, which would then require separate treatment with its own set of complications and costs over and above those of treating the underlying HIV infection.
3) Those who initiate treatment at a later stage in HIV disease progression are less likely to reach a full immunologic recovery (Okulicz et al., 2015; Asfaw et al., 2015; Maduna et al., 2015) – a finding, too, that is observed in our data set when looking at CD4 count recovery by CD4 count stratum at ART initiation.
   By delaying treatment, it would be expected that the patient would be at a higher general lifetime risk of contracting any opportunistic infection (Williams, Hargrove, & Humphrey, 2010).

While the health and economic impact of HIV/AIDS is well documented, and the clinical benefits of antiretroviral therapy have been quantified in terms of immune system recovery, decline in viral loads and the reduction in HIV related illness, the research on labour market outcomes comparing a treatment scenario to a theoretical no-treatment scenario in non-experimental settings is sparse. Essentially the research has looked at the effect of treatment while using the health state of each individual at the point of treatment uptake as the reference point for post treatment recovery. The effect of treatment is seen as the difference in the health state of an individual at the point of treatment uptake and the consecutive health states subsequent to treatment. While this does still

[2] We find that the general rollout of the universal test and treat model appears to have been gradually adopted across the company with no clear discontinuity in any specific year of the average CD4 count at ART initiation.

illustrate the benefits of treatment, the depiction of the full extent of the benefit will be understated since, mostly, the progression of the disease is stopped in its tracks at the point of ART uptake. These models do not directly account for the fact that there would be a likely deterioration in health from the 'date 0' had treatment been forgone. An interesting area of study, and one we later investigate, is to identify the probable health states of individuals had treatment been not been initiated at any stage of disease progression – the counterfactual to a treated case.

DATA AND METHODOLOGY

Our analysis aims to answer the following: How does treatment of HIV positive employees with triple combination ART affect health-related outcomes within a company's HIV-positive portion of their labour force?

This analysis entails, firstly, estimating the effect of treatment on the employee's health related outcomes, and secondly projecting employee health states had treatment been refused. In order to achieve this, we borrow from the techniques used by Habyarimana et al. (2010) with the use of a new sample, and in light of the use of more modern HIV treatment regimens. We also add insights to the literature on how treatment of HIV impacts the length and frequency of clinical visits in terms of both inpatient days and outpatient visits – a portion of in-house HIV treatment and care which makes up for about 50% of the total cost per treated patient at the firm around the treatment initiation date at the time of our study.

Data:

The data cover the period from January 2009 to December 2016 which have been aggregated to monthly measurements for each employee for clinical and absenteeism data; with annual updates occurring for individual level confounders such as marital status, age and job grade on the 1st of January each year. All clinical and absenteeism data is recorded on a per-incident basis with dates and observations attached to each incident for each person, tagged by a system generated person ID. The data has been aggregated to a monthly level and linked to the extrapolated HR data by person ID and month.

The treatment variable has been manipulated into a categorical variable which is able to isolate special treatment cases relatively well.
The categories for HIV specific treatment, which excludes other symptomatic treatments, are:

- Untreated: Those who are HIV positive but have never been on ART
- Treated: Those who are currently on ART
- Inactive: Those who have been on ART but are not currently on ART
- Own Doctor Treating: Those who have foregone company treatment and opted to get "treatment" elsewhere[3].

[3] Since this is a self-reported variable, we cannot be sure of the legitimacy of the external treatment. In some cases, patients seeking traditional or cultural treatment may report to have opted for external "treatment" not in the form of ART.

The data contain 12 035 unique person observations, of which 5 072 (42,14%) of the people remain in the panel across the entire observational period. 75% remain in the study window for at least 47 months. The average employee count over the years is roughly 8 900. We observe 1962 separate cases of HIV positive employees either from the start of the window or from some point during the window of observation. 47% of the people who have been or will be diagnosed with HIV at any point during the study remain employed at the company for the full time period under study, with 75% of this sample remaining employed for at least 54 months. Over the 8 year window there were 583 879 illness related absenteeism incidents, 184 684 outpatient visits and 11 212 inpatient days recorded. There were 1 243 separate treatment initiations across 1 028 HIV positive employees who have initiated treatment during the sample period, of which 880 are first time treatment initiations. A group of 932 patients has reported complete adherence to their treatment regimen during their time at the company (which may extend beyond the sampling window). Treatment initiations across the sample window are summarized in figure 3.

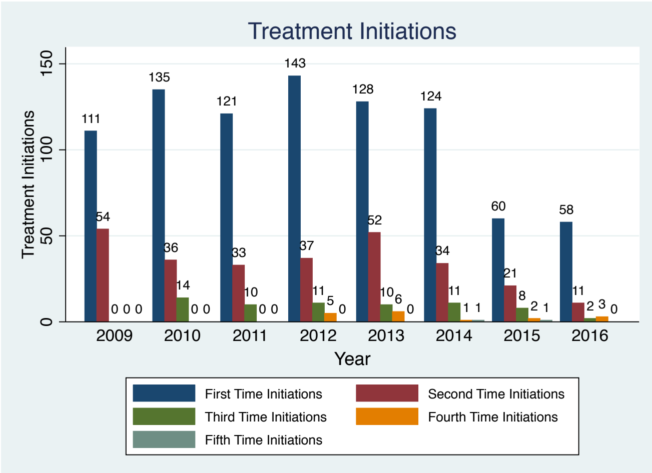


Figure 1

With the introduction of the HIV treatment programme, policy stipulated CD4 counts were to be tested on a quarterly basis for those who had enrolled in the programme. An average duration between clinical tests of 3.6 months was observed in 2009 with a general upward trend in the length between tests to 5.7 months in 2016, as the company deemed the quarterly frequency of tests to be excessive for clinical monitoring purposes in most cases.
For certain estimations, we use linear interpolation in order to populate missing CD4 counts between tests.
The treatment initiation date and the HIV programme enrollment date need not necessarily coincide - a median gap of enrollment in the HIV programme (which includes clinical tests but not necessarily treatment) to first treatment initiation of 13 months is seen in the data.

The distribution of CD4 counts at ART initiation has slowly shifted rightwards over the years, due mainly to the steady adoption of more relaxed treatment eligibility criteria.
The gradual upward shift in the proportion of ART initiations at higher CD4 counts is documented in figure 4. Group B, the percentage of those initiated on ART prior to progressing to a WHO defined advanced immunosuppression (< 200

cells/mm³), has increased drastically since 2013, while those in group A, those initiated on treatment only after progressing to an advanced level of immunosuppression has dropped to about 13%. The greatest area of growth in the proportion of initiations is in the CD4 stratum of >600 cells/mm³ at first-time ART initiation, while those initiated at a CD4 count of less than 50 has fallen to 0%.

There has been a paradoxical increase in the HIV prevalence rate of the tested at the company since the implementation of the programme. This is due in part to higher test rates, resulting in more accurate estimations of the population prevalence rate over the years as well as being an artifact of treatment itself – fewer deaths with somewhat stable new infection rates result in higher prevalence rates. The HIV prevalence rate as of 2016 is 16,53%[4].

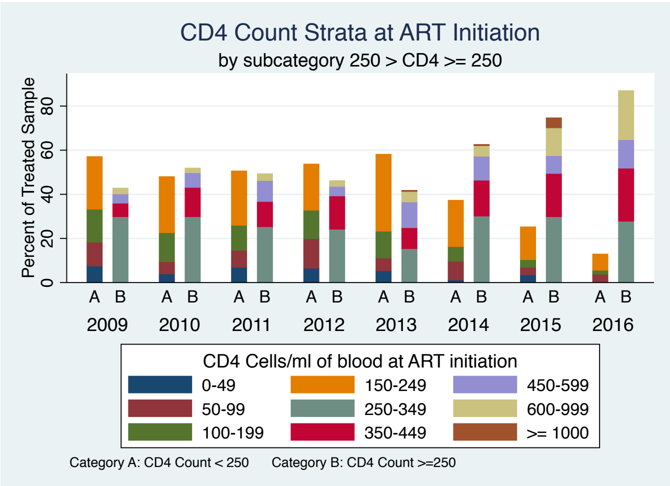


Figure 2

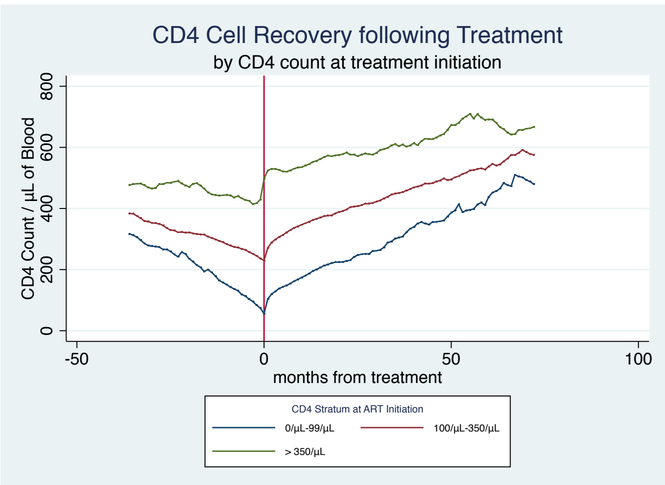


Figure 3

Although the statistical causality of the relationship between HIV treatment and TB incidence at the company is yet to be established, the remarkable decline, by more than 2 thirds, of newly diagnosed TB cases is somewhat of a medical marvel warranting further research into the topic.

Although not central to the paper, we depict briefly the relationship between CD4 counts and incident TB, what could be loosely be inferred as a proxy for all opportunistic infections, in the HIV positive sample[5].

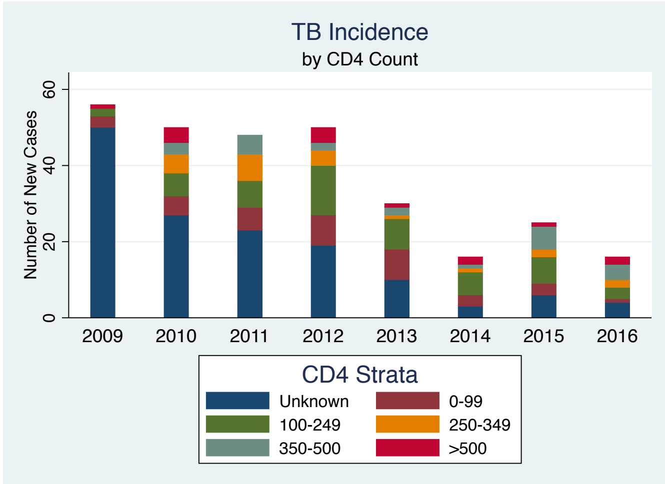


Figure 4

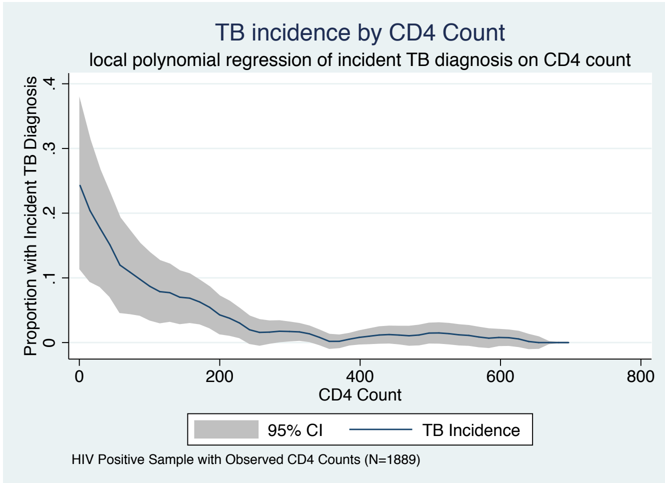


Figure 5[5]

[4] Test rates and prevalence rates by year may be found in the appendix

[5] We restrict the regression to HIV positive employees with observed CD4 counts. This does introduce a form of bias in that those HIV-positive patients diagnosed with TB will, as a matter of policy, be urged to test their CD4 count, while those without a recent TB diagnosis are not likely to be presented with the same urgency to test CD4 counts. We would find higher CD4 test rates across samples with active TB than across samples without active TB. Despite the possible overall bias, we would expect the general trend to remain regardless of the possible introduction of selection controls.

ESTIMATION AND MODELING

The Problem of Estimating Treatment Effects in the Presence of Strong Selection into the programme as well as HIV treatment uptake:

At the heart of this analysis lies a company treatment programme in which treatment uptake, although now encouraged at every stage of HIV progression, is an entirely voluntary decision for the patient.
Estimating the treatment effect of the programme is also severely complicated by aspects internal to the company, such as the fact that, originally, eligibility for treatment was determined by either a threshold CD4 count which is a strong determinant of the progression of the disease, or for those who had already contracted a severe opportunistic infection such as TB. This resulted in, generally, asymptomatic HIV-positive employees making up the majority of the untreated group, while the treated group would be those who were already presenting with symptoms of immunosuppression around the treatment initiation date, thus creating a substantial treatment uptake selection bias. Beyond the selection into treatment from the company's side, in the earlier years of the programme, and even after the treatment eligibility was extended to all HIV positive employees regardless of disease progression, it would not be unreasonable to assume that the potential adverse side effects of treatment, the introduction of potentially time consuming daily routines, or the pure stigma of HIV could result in HIV positive people in the very early stages of disease progression delaying treatment until the perceived health benefits of treatment outweighed the perceived costs. We speculate that selection bias thus arises from both the company's side as well as the patients' sides. Even once possible selection bias is controlled for based on observables such as pre-treatment CD4 counts, viral loads and time since diagnosis which would directly affect health outcomes, the decision to initiate treatment from the patient's side at any stage would still, essentially, be a behavioural phenomenon which is difficult to directly observe. With this possible behavioural link, and as with the correlation between lifestyle choices and HIV infection (Robles et al., 1994) or willingness to test for HIV (Bärnighausen, Bor, Wandira-Kazibwe, & Canning, 2011), those who choose not to enroll into the programme could be the ones with lifestyle choices that also adversely affect health outcomes unrelated to HIV infection, and, consequently, could also affect our outcome variables.
We look to solve part of the problem of selection by restricting our analysis to only those patients who have opted for treatment at any stage during our 8-year sample window, while disregarding those HIV positive employees who are yet to initiate treatment. The panel structure of our data allows us to observe, across time, any patient meeting our sample selection criteria in the months prior to and following treatment uptake – a untreated and subsequently treated state. This will allow us to estimate the dynamics of health and health recovery prior to and following treatment uptake
It is worth noting, however, it would be expected that there would be disparate effects of treatment on viral loads, CD4 counts, days absent, as well as inpatient days and outpatient visits contingent on the stage of disease progression at treatment uptake (Asfaw et al., 2015). On the two extreme ends of the spectrum, those who initiate treatment at a late stage are likely to see a drastic and prompt recovery from their current health state (barring the paradoxical phenomenon known as immune reconstitution inflammatory syndrome (IRIS) in very late stage AIDS patients), yet those who initiate treatment at a completely asymptomatic stage are likely to benefit only from the fact that there will not be a decline in their already adequate health state following treatment (Sharma & Soneja, 2011; Maduna et al., 2015).
In light of this, we will discuss two possible concepts around the effect of treatment in our analysis.

1) The average effect of treatment as it stands at the company for the pool of people who initiated ART at the observed levels of disease progression.
2) The average difference between the observed treated states of the individuals following treatment and the modeled state that would have been observed had treatment been forgone.

As a first pass approach, we adopted matching techniques in order to identify this counterfactual - finding suitably similar people based on conditions around treatment uptake, disease progression, or time from infection, while controlling for other factors, and differing only in that of treatment status. In this way, we could use the untreated person as a theoretical reference case for what would have happened to the treated patient in the absence of treatment. However, we find, especially in this sample, a distinct lack of untreated patients who could be identified as being a match to a treated patient in the treated patient's treatment initiation month and then subsequently staying untreated for a long enough period of time to observe the true difference between treated and untreated cases. It is also a possibility that the patients who are assign themselves to the untreated control may never progress to a point of immunosuppression in which they would justify treatment during our sample period. The complexity of this is brought about mainly as an artifact of the disparity observed in time to disease progression across patients. The matched untreated controls will generally opt for treatment the moment a severe enough random health shock is experienced - otherwise all that is observed over time in our control is an, essentially, asymptomatic HIV-positive patient with a similar CD4 count and time since diagnosis as the treated patient, as measured at the treated patient's ART initiation month.
As an intuitive method, we opt, instead, to attempt to model theoretical disease progression in the absence of treatment for each treated patient as of each patient's treatment date based on observed disease progression prior to ART uptake, as well as on external clinical studies on disease progression based on CD4 counts. This would allow for a counterfactual to the treated case to be estimated, while explicitly modeling the fact that, had treatment been refused, the health state of the treated patient would likely decline from that of the observed health state as is measured at the point of treatment initiation.

Adopted Modeling Approach:

Following in the style of Habyarimana et al. (Habyarimana, Mbakile, & Pop-Eleches, 2010), the average impact of the provision of antiretroviral therapy on the number of sick days taken by an HIV positive employee in any given month will be modeled primarily on the number of months prior to and following treatment uptake, while controlling for both month and year time trends independent of the controls for the months from treatment. We also include an array of controls for observable individual level confounders.

Additionally we look to control for unobservable individual characteristics and sample attrition using a combination of fixed effects models and inverse probability weightings of sample observations in separate models.

The estimation differs in that of the estimation of Habyarimana et. al. as those who are not in the pool of people who have been or will be diagnosed with HIV are left out of the estimation (ie. We look only at those who are consistently HIV positive across the sample window). This will, in effect, show the correlations between sick days, as well as inpatient and outpatient days and the dependent variables for only the HIV positive pool in our dataset, which is our primary sample of interest. Part of the distinctiveness of this data set is that it benefits from a suitably large cohort that suits our selection criteria.

We model our outcomes on the general estimations which will be determined by the following functional form, estimated by OLS:

1) $$y_{it} = \beta_0 + \sum_j \gamma_j\, months\_from\_ART\_start_{it}^j + \sum_k \beta_k\, person_{control}{}_{it}^k + \sum_h \tau_h\, month_{it}^h + \sum_p \lambda_p\, year_{it}^p + \varepsilon_{it}$$

where:
estimation 1)
y_{it} = number of Illness related absenteeism events per month,

estimation 2)
y_{it} = number of inpatient days per month

estimation 3)
y_{it} = number of outpatient visits per month

and common to all estimations:
$months\_from\_ART\_start_{it}^j$ = indicator variable for each month, $j \in [-11; 12]$, from the treatment initiation month for employee i at time t, where the base month is -12 months from treatment uptake.

$person\_control_{it}^k$ = each person control, k, for employee i at time t.
k = {gender, age, age^2, marital status, employment band}

$month_{it}^h$ = indicator variable for each month of the year, $h \in [2; 12]$, for employee i at time t, where the base month has been set to January.

$year_{it}^p$ = indicator variable for each year, $p \in [2010; 2016]$, for employee i at time t, where the base year has been set to 2009.

ε_{it} = the error in estimation for person i at time t.

Sample attrition

Beyond the obvious sample selection problems brought about by such voluntary treatment uptake programmes, we explore the possible attrition of our sample of those who have opted for treatment (those who have started treatment and subsequently discontinued treatment after any given number of months) which could potentially introduce yet another source of bias.
The concern that attrition in our sample, and the associated missing data from lost participants, could be fundamentally correlated with absenteeism, inpatient days and outpatient visits creates a scenario in which it is worth looking at separation patterns within the sample. Conceptually, the sickest people are most likely to fall out of the sample due to ill health retirements or death, and had these people remained in the sample, would have likely had a heavy weight on the average observed values of the outcome variables. Looked at from another point of view, the healthiest individuals following treatment, for which ART has been wholly effective, may be of the opinion that they need not continue treatment. These patients could conceptually bias the results of our estimations in the opposite direction as previously discussed by dropping out of our sample of treated patients. Ideally, we seek a scenario in which those contributing to the attrition in our sample are dropping out randomly and not due to some relationship with their health state at the point of separation.

In figure 5, we consider the possible impact sample attrition could have on our analysis - not as a measure of true treatment attrition rates at the company. The attrition patterns are depicted purely for those who are used in the sample in the identification of our regression model.
The attrition rates shown in figure 5 are split into groups of those who have stopped first round treatment for any reason, be it non-adherence, adverse reactions etc., and those who have fallen out of the sample/left the company due to ill-health-retirements or death. These attrition rates are geared directly at first time initiations, regardless of whether treatment is then started at a later date in the form of a second, third or fourth round initiation[6].

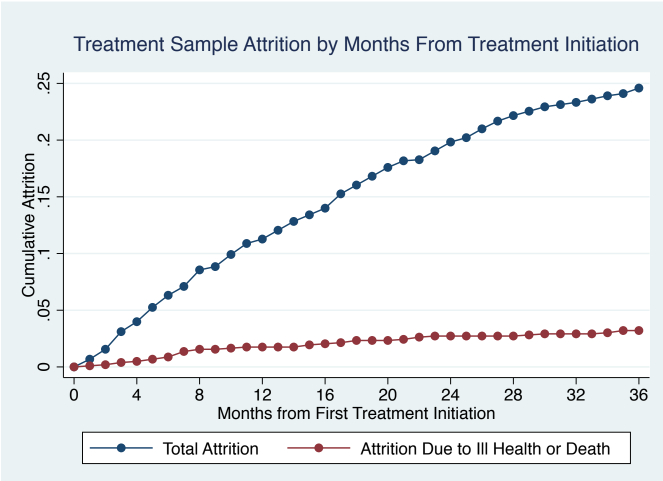


Figure 6

The total attrition, attributable to any scenario, of participants undertaking first-time treatment in this sample is around one in four patients after 3 years, with a steady decline in the rate of attrition the greater the length of time from first treatment

[6] We look purely at first round initiations due to the fact that there is no reliable way to have a single measure of the months from treatment initiation in a case where someone has stopped treatment and then started again. A person who stops first time treatment after five months and then starts treatment for a second time 10 months after starting the first time could be considered as a patient both +5 months from treatment and -5 months from treatment.

uptake[7]. The total attrition rates are similar to those observed in Habyarimana, yet those due to ill health or death seem to be considerably lower, which lies at 4% as opposed to 12% after 3 years The result provides some relief that this may not be as large of a concern in causing significant sample attrition bias.

On the other spectrum of the discussion on attrition, we do not allow for replenishment within the sample. Those who would be replenishing the sample are those who would be initiating second, third or fourth round treatment. The interaction between months from first treatment and that of subsequent treatment initiations and the related influence on health status is hard to theorize in this scenario since the outcome is based largely on the situations surrounding treatment termination. We assume the "best case" scenario by limiting our sample to those of whom have reported complete adherence to treatment for at least 12 months following first round treatment initiation. Adding to the former issue is that any individual stopping treatment for any reason and then starting treatment at a later date can be considered to be both a positive number of months from first treatment and a negative number of months from second treatment, where both first and second time treatment initiations are considered as date 0 (treatment initiation month). We thus exclude any patients without a reported 100% adherence to treatment[8].

The estimations describing the underlying relationship between months from treatment uptake and our outcome variables, monthly sick days, inpatient days, and outpatient visits, are run using an array of functional forms for comparison of the robustness of the estimates.

Due to the fact that the decision to seek treatment could be highly correlated with individual level behavioural characteristics, such as levels of risk aversion, which may also impact absenteeism unrelated to HIV-related events, an individual fixed effects model is fitted to control for unobservable characteristics, over and above the controls for the observable time variant factors.
In the spirit of Habyarimana et. al. we break the analysis down into one in which an unbalanced panel is used with and without fixed effects, a balanced panel and, finally, a model using inverse probability weights in order to control for sample attrition and the possible bias this may introduce.

In our use of the inverse probability weighting of sample observations, where an assumption is drawn that those who stop treatment for any reason, whether it may be death, ill health (and separation from the company) or treatment non-adherence, do not do so randomly and that attrition is fundamentally correlated with the outcome variables. The inverse probability weighting will, to some extent, control for any attrition bias that may be present. We are assuming that when estimates are conditioned on the inverse of the probability, based on observables, of someone dropping out of the sample, any attrition and related missing data will be a random occurrence, such that the ignorability assumption is satisfied (Weuve et al., 2012; Wooldridge, 2002, 2007).

The probability weights for each employee for the IPW estimation are calculated as of the date of first treatment initiation based on the probability of the programme participant leaving the treatment programme within 12 months of the first treatment start date. Those who have a high probability of leaving the programme based on observable characteristics at date 0 and yet stay in the programme for at least 12 months are given a higher weight in order to compensate, in a way, for the attrition in the sample that can be attributed to the factors included in the probability model and that would otherwise bias the results.
The probability model does not look at the factors influencing the participant having been observed in the months prior to treatment – this would be purely based on how long the participant had worked at the company prior to treatment uptake, and not necessarily as a result of the magnitudes of any of the outcome variables, since *all* programme participants would have observations of both independent and outcome variables if they had been at the company for at least 12 months prior to treatment initiation.

We specify a likely probability model for use in calculating our probability weights, based on patients' health states at treatment initiation and using the employees' genders, marital statuses, job grades, a quadratic in age and maximum tenure at the company as of treatment initiation, as well as the month and year of treatment initiation and starting CD4 count. The number of times a person has left and rejoined the company seems a likely descriptor of the probability of leaving the sample at the moment of treatment uptake, however, there is a perfect correlation between leaving the sample within 12 months and having previously left the company, we are thus required to exclude this descriptor from the estimation of the probability weights.
The exact specification, estimated at the point of treatment uptake, using maximum likelihood, is as follows:

$$logit(p_i) = \beta_0 + \sum_k \beta_k\, person\_control_i^k + \sum_h \tau_h\, month_i^h + \sum_p \lambda_p\, year_i^p$$

where:

p_i = the probability of leaving the sample within 12 months of treatment uptake

$person\_control_i^k$ = the vector of individual level controls for person i,
{k = Age, Age^2, Gender, Marital_Status, CD4_Count_at_ART_Start, CD4_Count_at_ART_Start^2 Job_Grade, Maximum_Tenure_at_Company, Maximum_Tenure_at_Company^2}

$month_i^h$ = indicator dummy for the month in which treatment was initiated for patient i
h ∊ [2;12]

$year_i^p$ = the year in which treatment was initiated for each patient i
p ∊ [2010;2016]

ε_i = the error in estimation for person i at time t.

Results and Discussion:

[7] Separations for the total company, both HIV negative as well as positive sits at around 15% after 3 years.

[8] There will almost always be a disparity between reported and actual adherence. Despite this, a recorded treatment termination due to patient non-adherence is what would trigger an official failure of a first time treatment initiation in our analysis.

The results of the various estimations are presented in table 1 for the impact of ART on sick days. Although also discussed in this section, the estimations for the impact of ART on inpatient days and outpatient visits may be found in the appendix to this paper.

pronounced effects in only the magnitude of changes for lower CD4 strata at ART initiation. What we would find in plotting these changes in health states around treatment is 3 nested

curves, with the group of the lowest CD4 counts displaying the greatest magnitudes in changes.

Table 1

	Sick Days	Sick Days (Unbalanced Sample)	Sick Days (Balanced Sample)	Sick Days (IPW Estimation)
Fixed Effects	N	Y	Y	Y
11 months before treatment	-0.00996	-0.000973	0.0994	0.114
10 months before treatment	-0.0290	-0.0118	0.103	0.0851
9 months before treatment	0.0744	0.0980	0.175	0.107
8 months before treatment	0.131	0.160	0.0963	0.0446
7 months before treatment	0.187	0.228	0.191	0.221
6 months before treatment	0.154	0.204	0.212	0.218
5 months before treatment	0.415*	0.473**	0.566**	0.505*
4 months before treatment	0.597***	0.650***	0.842***	0.787***
3 months before treatment	0.687***	0.752***	1.081***	1.168***
2 months before treatment	1.062***	1.142***	1.527***	1.889***
1 month before treatment	1.816***	1.913***	2.326***	2.574***
Treatment initiation month	4.757***	4.873***	5.308***	5.797***
1 month after treatment	3.175***	3.290***	3.902***	4.455***
2 months after treatment	1.898***	2.025***	2.299***	2.408***
3 months after treatment	1.277***	1.413***	1.829***	1.752***
4 months after treatment	1.068***	1.215***	1.593***	1.412***
5 months after treatment	0.915***	1.072***	1.333***	1.191***
6 months after treatment	0.795***	0.961***	1.110***	1.020**
7 months after treatment	0.505**	0.682**	0.926***	0.718
8 months after treatment	0.634***	0.822***	1.145***	1.024**
9 months after treatment	0.454**	0.653**	0.983***	0.857*
10 months after treatment	0.218	0.422	0.449	0.542
11 months after treatment	0.328	0.537*	0.822**	0.629
12 months after treatment	0.285	0.499*	0.758**	0.582
Observations	20,929	20,929	13,539	18,567
R-squared		0.063	0.073	0.094
Number of unique id	1,001	1,001	500	799
*** p<0.01, ** p<0.05, * p<0.1				

Notes:
Each of the above specifications contain controls for month and year time effects distinct from the variable measuring time from treatment uptake, other controls include a quadratic of the patient's age, and in the case of the non-fixed effects regression, gender, job grade and marital status.
Similar regressions covering the analysis of inpatient days and outpatient visits may be found in the appendix to this paper.

The coefficients of the dummy variables of months around treatment initiation show, prior to the initiation date, the estimated disease progression in terms of its effect on sick days, inpatient days and outpatient visits, and following treatment, the estimated effect of ART over time while controlling for other factors which may influence the outcome variables of interest. Statistically insignificant coefficients of any particular month from treatment suggest no marginal impact above the reference month, 12 months before treatment, which, given the progression of CD4 counts previously illustrated we could consider to be a point in which patients would present with few symptoms of immunosuppression. We observe statistically insignificant impacts of most of the months building up to treatment initiation date which coincides with the asymptomatic period of disease progression. In the four to five months before treatment initiation, we observe a noticeable increase in the number of sick days, inpatient days and outpatient visits. All three of these variables peak in the month of treatment initiation, again confirming the large degree of selection into treatment. The marginal impact of treatment in the months following treatment initiation is drastic, with a reduction in sick days, on average, of one day per month, additively, for each of the first three months following treatment, with further yet less pronounced improvements in health for up to 10 months following treatment initiation.
Delving further into the trends revealed by this regression, we present in the appendix the results of the above regression, stratified into 3 groups by CD4 count at ART initiation. We find that the broad patterns around the treatment initiation date are remarkably similar across the strata yet with more

After progressively rising from statistically insignificant values in the 4 months prior to initiation, inpatient days fall from a peak average across the sample of 0,6 days in the treatment month to statistically insignificant values after 4 months.
Outpatient visits, the dynamics of which are slightly different to those of sick leave and inpatient days since outpatient visits can, at any stage, be attributable to non-illness related visits (such as renewal of scripts for chronic medication), reach a peak of 3,5 visits per month in the treatment month, however remain at just under one visit per month on average even after full recovery. This can presumably be ascribed to checkups relating to patient treatment adherence, general health, and Although the above specifications illustrate the effect of treatment on those who opt into the programme, they do not, in any way, depict what would have been the outcome had the patient, at the time of treatment uptake, been refused, or, themselves, refused treatment. We have a fundamental issue in that the estimations of treatment effects on the treated sample is exactly that – estimations for a treated sample. The greatest area of interest, however, when modeling the full benefit of treatment is to depict the outcomes in a case where treatment was not given at the time each person in our sample was, in the reality, initiated on ART.
We look to address this issue by modeling a plausible path of disease progression, in the absence of treatment, in the form of CD4 decline for each patient at the point of treatment initiation and beyond, and subsequently mapping clinical disease progression to each of our outcome variables.

CD4 modeling and Counterfactual estimates for Sick days in the absence of treatment:

The fundamental issue faced when looking for treatment effects in this analysis is the lack of a suitable statistical counterfactual. We attempt to model the theoretical no-treatment scenario for each treated patient by mapping disease progression, in the decline of CD4 cells in the absence of treatment, to absenteeism data. We will, in this way, be able to plot out a probable time path of sick days in the absence of treatment following the actual observed treatment date of each patient. This method does not come without its issues – chief among which is the fact that we are essentially attempting to fit a model using estimates we are assuming to have sufficient external validity to extrapolate values of CD4 decline into our data set, and again using a best guess at the relationship between CD4 counts and sick days. CD4 counts in level form, although highly correlated with absenteeism can only explain roughly 6% of the variation in sick days. Since the CD4 count, beyond the month and year controls, is the only variable we are predicting a time path for in our model, we are assuming that this is the primary mechanism driving the dynamics of the incidence of sick days – a somewhat unrealistic assumption. In spite of the troubles of using this technique, it nonetheless provides plausible estimates of the time to death from what we will call the theoretical "refusal of treatment date", date 0. Providing encouragement for this technique is the fact that the estimates are not out of line from what previous literature investigating health deterioration of HIV positive patients would predict. Further to this, the estimates of the progression of sick days would seem to be, if anything, understated when compared to the time path of in-sample sick days prior to treatment.

In order to model CD4 count decline from the date of treatment in the theoretical case of the person not having opted to initiate treatment, we use the nearest CD4 count prior to treatment in order to estimate the CD4 count at date 0 should the last CD4 test date not be in the same month of treatment uptake. Generally speaking, and in the earlier years of the programme, CD4 counts were the trigger for treatment initiation, or, at the very least, testing of CD4 counts had and continues to have a high enough correlation with treatment initiation as to have programme participants' CD4 count tests in the month of treatment, or in the one or two months prior to treatment initiation. This relationship has bidirectional causality. Our strategy of interpolation is thus expected to yield more or less accurate estimates of CD4 counts in the month of treatment. Nonetheless if there is a treatment initiation date prior to which, the nearest CD4 count observation is greater than 3 months from the date, we exclude the estimate from our interpolation strategy. The relationship between disease progression and health outcomes such as viral loads and CD4 counts is well documented in the medical literature (Dunn, 2006; Hogg, 2001; Holmes et al., 2006) with evidence pointing to an average decline of 12 CD4 cells per month for those with CD4 counts below 300 cells in the absence of treatment (Jaffar et al., 1997; Morgan et al., 2002). As depicted in table 2, the average CD4 counts for our sample, at ART initiation prior to 2015, falls within an acceptable range of the 300 cells/mm^3 baseline for us, for the most part, to expect the sample to adhere to similar dynamics of CD4 decline. Although the 2016 initiations fall far from the 300 cell mark, the vast majority of these patients will not make it into our sample, since we previously stipulated the requirement that there be 12 months' worth of observations following treatment uptake.

Mean CD4 Count at ART Initiation by year

2009	209.3503	2013	237.0396
2010	241.9107	2014	295.0774
2011	214.9615	2015	387.9898
2012	232.1667	2016	428.4242

Table 2

The relationship between disease progression and sick days/ productivity has been documented in 'The Impact of HIV/AIDS on Labour Productivity in Kenya' (Fox, et al., 2004). Although we could use the aforementioned results directly in the modeling of sick days in the absence of treatment, we opt for the more indirect approach of first parameterizing a model based on the relationship between our outcome variables and patients' CD4 counts, following which, we extrapolate CD4 counts beyond the treatment date using the expected CD4 decline in the absence of treatment, and subsequently employ the aforementioned models to obtain a simulated counterfactual. We expect this method to have reasonable internal validity. This approach also allows us to maintain consistency in our modeling of sick days, inpatient days and outpatient visits.
This is not an ideal scenario since we would prefer to model the total treatment effect based on an 'observed' counterfactual, however this is generally not possible.
The span of the asymptomatic period of the disease can differ drastically between patients. To observe a possible alternate to a treated patient we would require that an untreated patient be similar in all attributes other than their decision to be treated. These patients in this setting are few and far between, since there is such high treatment uptake. Beyond this, the measurement of CD4 counts, one of the best predictors of disease progression as well as being the best matching variable in this dataset, is also highly correlated with treatment. We find that the untreated sample has fewer CD4 measurements in which to use to match to a treated patient.

Despite the challenges of modeling a counterfactual by employing the method that follows, our model does yield highly plausible results and is generally in line with what we would expect in terms of survival rates as well as the general decline in health and associated increase in monthly sick days over time (Bor, Tanser, Newell, & Bärnighausen, 2012; Charalambous et al., 2007, 2007; Fox et al., 2004; Meyer-Rath et al., 2012; Mocroft et al., 2007; Sonnenberg et al., 2011).

We diverge from the modeling of a panel fixed effects regression for the estimation of the relationship between CD4 counts and sick days, inpatient days and outpatient visits since the coefficients of the time invariant factors are required for the prediction of the outcome variables which the fixed effects model would difference out. We control for as many relevant observables as possible, however, a limitation of this is that regardless of how comprehensive the variable list is we cannot control for unobserved behavioral mechanisms that may be influencing the number of sick days taken.

We use a cubic function in CD4 counts[9], a quadratic in age and consecutive months worked, indicator variables for the year and, separately, month of the year, gender, marital status, and worker band.
The results of the linear regression depicting the relationship between that of sick days taken per month and CD4 counts are presented below (Table 3). We use the entire HIV positive sample in this estimation, which is why there is a jump in the number of unique ID observations.

Table 3

	CD4_Sick_days	Standard errors *** p<0.01, ** p<0.05, * p<0.1
CD4_Count	-0.0194***	
CD4_Count2	2.18e-05***	
CD4_Count3	-7.13e-09***	
month = 2	-0.0141	
month = 3	0.124	
month = 4	0.0706	
month = 5	0.269*	
month = 6	0.306*	
month = 7	0.731***	
month = 8	0.359**	
month = 9	0.478***	
month = 10	0.447***	
month = 11	0.212	
month = 12	-0.170	
year = 2010	0.0697	
year = 2011	0.123	
year = 2012	0.523***	
year = 2013	0.705***	
year = 2014	0.396***	
year = 2015	0.424***	
year = 2016	0.274*	
Gender = Male	-0.123	
Marital Status = Single	-0.00847	
Current Age	-0.00632	
Current Age^2	1.10e-05	
Constant	5.942***	
Observations	18,054	
Number of id	2,164	

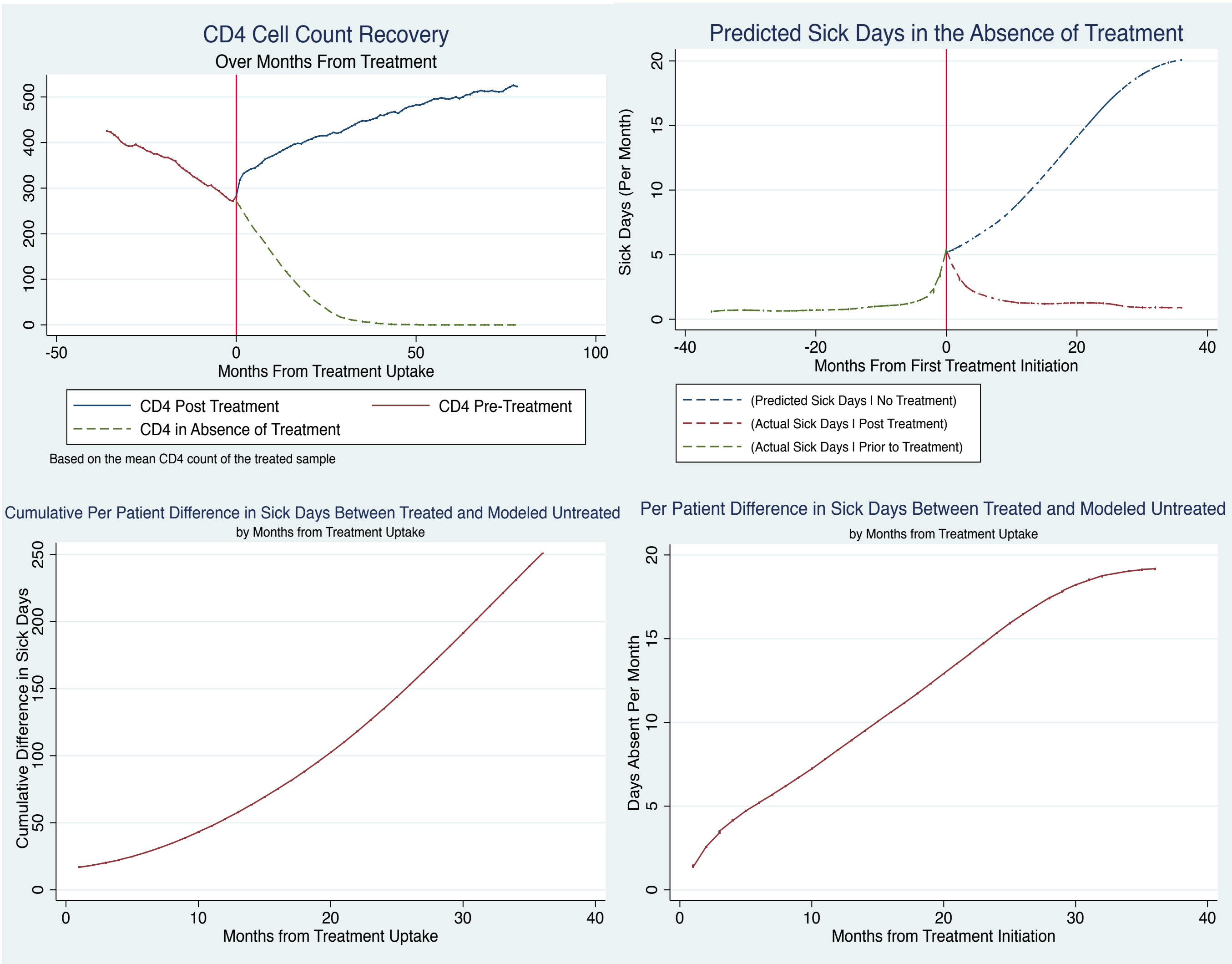
The above results are particularly sensitive to the use of imputed values of CD4 counts, most probably due to the fact that the measurement of CD4 counts is likely correlated with CD4 counts themselves. The measured CD4 test observations, or lack thereof, in a standard regression would result in a regression based purely on those patients who needed a CD4 count test in the first place in that particular month. Although there are a multitude of factors that could prompt such a test, of which some are mostly unrelated to the observed health state of the individual, other reasons that would trigger medical staff to test CD4 counts in HIV positive patients would be the possible diagnosis of opportunistic infections. Thus, missing CD4 counts in our data would not be missing at random. To address this, for the most part, we use linear interpolation to fill missing observations between any two CD4 tests. As with papers by Habyarimana et al. (2010) and Meyer-Rath et al. (2015) we assume the linear prediction of CD4 counts between observations is a more or less accurate approach in dealing with missing values. We

We account for weekends and public holidays in the analysis, giving the maximum possible number of sick days as the number of work days in each month. We assume death occurs at CD4 counts less than 30 cells, resulting in 100% sick leave. The average CD4 count at the treatment date for those initiating treatment is 253 cells/mm^3 with a standard deviation of 160 cells/mm^3. We find that, had each patient not initiated treatment, 75% of the group initiated on treatment would be dead after 25 months, and the entire HIV positive workforce would be dead within 51 months had they forgone treatment. This is in contrast to the Habyarimana paper in which it is calculated that all enrolled workers would be dead within 26 months from date 0 had treatment been foregone. We can attribute this difference to earlier enrollment in the programme in the Anglo case.

Across figures 7-10, we depict, graphically, the results of our modeled counterfactual outcomes. We include representations of the counterfactual in CD4 decline in the absence of treatment as well as the implied increase in the amount of sick leave taken relative to the case in which treatment was initiated.

The full table of results to be discussed may be found in the appendix.

[9] Missing values are imputed using linear interpolation between any two consecutive measurements of CD4 counts for each person. On average there are between 3 and 4 imputed values for CD4 counts in the months between any two actual measurements of CD4 counts.



We calculate the treatment effect for the individuals for each consecutive month following treatment uptake as the difference between the actual sick days post treatment and the modeled sick days for that individual based on CD4 decline had treatment been refused. Figure 7 shows the changing treatment effect as the distance from treatment grows. The greatest change in the treatment effect is in the first few months following the treatment date where those initiated on treatment display a drastic recovery while the modeled no-treatment case shows a slow and steady decline in health. The predicted treatment effect stabilizes as the treated reach pre-illness health states while the modeled non-treated sample continues its steady decline. Our modeled treatment effect is calculated in such a way that it displays the dynamic nature of disease progression in a way that most other papers omit.
We can depict both the gradual recovery of the treated relative to the treatment date as well as the gradual decline in health of the modeled untreated and the widening gap between the two states over time from date 0.

The total treatment effect over n months is calculated as the cumulative difference between the treated and untreated states for each person up to month n – Essentially the integral of the graph of the per patient difference in sick days between treated and untreated cases. After 36 months from the true treatment date, we estimate that, had treatment been forgone, the untreated patient would have taken a total of roughly 250 more sick days during this period.

Limitations

Although the results presented paint a strong picture that are of interest to policy makers, the limitations introduced by the non-random selection of the data in the form of looking purely at those with full adherence would suggest a difference in true overall policy outcomes. Disregarding patients who have disrupted treatment at any stage is considered to be weakness of this analysis. Further work could look into the outcomes for all adherence groups.
Non-random missing instances of missing data that could, intuitively, be correlated with health outcomes introduces another possible source of bias – those with missing CD4 count measurements could be those who feel healthy enough to justify missing a health checkup to themselves, or otherwise could include those too sick to make the journey to a clinic.

Conclusion

The patterns of absenteeism around the date of ART initiation depict the rapid progression of HIV and related immunosuppression in the months building up to treatment dates. The increase in absenteeism, inpatient days and outpatient visits can all be related to the decline in CD4 counts which, across all CD4 strata, reach their minimum in the month that treatment is initiated.
Although CD4 counts recover post treatment across all CD4 strata, those who initiate treatment at lower absolute CD4

counts have persistently lower CD4 counts for many years following treatment.
Treatment programmes would benefit from a faster rollout of 'universal test and treat' strategies in this regard.

We find that the reduction in sick days, on average, for those initiating treatment is most drastic in the first 3 months following ART initiation. Compared to the incidence of absenteeism attributable to the distance from treatment initiation, patients take on average 3 fewer sick days per month than in the month of treatment uptake. After the first year on ART, patients take on average 4,5 fewer sick days per month than in that of the month of treatment uptake.

Patterns of inpatient days around treatment uptake display similar dynamics to that of absenteeism in that they rise steeply in the 4 months prior to treatment uptake and drop off rapidly in the first 3 months following treatment. The demand for healthcare resources at the company can be well approximated by these figures. As with the relationship between TB incidence and CD4 counts, the majority of inpatient visits can be attributed to patients initiating ART at more advanced stages of immunosuppression.

Outpatient visits, unsurprisingly, peak in the month of treatment initiation before falling and then stabilizing at around 1 visit per month. These visits can be attributed to follow ups and script renewals.

The benefits of antiretroviral therapy for HIV positive individuals in terms of decreased mortality, morbidity as well as an increase in quality of life has broader effects for company efficiency. The management of employee health can be and should be equated to any other programme aimed at increasing the efficiency of company operations.
The effective management of health, of which the management of HIV plays an increasing role in Southern Africa, has net benefits for employee productivity.
The cost of treatment and care for HIV patients has dropped so substantially in the last decade that the strategy of providing this medication free of charge to employees could be cost saving, even in the absence of much deserved public-sector subsidies.

The benefits for large companies, at a micro level, of the implementation of HIV treatment and care strategies, if allowed to permeate to small and medium enterprises, too, could potentially have spillover effects that extend far into the macro economy.

Antiretroviral therapy has given HIV positive people a second chance at living healthy lives. Beyond the benefit to the individuals, businesses share in the value created by the provision of this life changing medication to their workforces. Businesses must learn to be proactive in fighting a disease whereupon its demise will be completely determined by both the public and private sectors' response to an epidemic which could otherwise cripple economies over many generations.

APPENDIX

Inpatient Days [10]

VARIABLES	total_inpatient_days Unbalanced No Fixed Effects	total_inpatient_days Unbalanced Sample	total_inpatient_days Balanced Sample	total_inpatient_days IPW Estimation
				*** p<0.01, ** p<0.05, * p<0.1
Fixed Effects	**N**	**Y**	**Y**	**Y**
11 months before treatment	0.0564	0.0573	-0.00870	0.0194
10 months before treatment	0.0349	0.0372*	0.00577	0.0180
9 months before treatment	0.0163	0.0205	0.0116	0.00542
8 months before treatment	0.0221	0.0276	-0.0108	-0.00214
7 months before treatment	0.0310	0.0395	-0.0114	-0.0114
6 months before treatment	0.0278	0.0378	-0.00802	-0.0127
5 months before treatment	0.00871	0.0220	0.00115	-0.00941
4 months before treatment	0.105*	0.121***	0.114**	0.0860
3 months before treatment	0.156***	0.175***	0.244***	0.182**
2 months before treatment	0.157***	0.180***	0.218***	0.181***
1 month before treatment	0.308***	0.337***	0.351***	0.295***
Treatment initiation month	0.576***	0.597***	0.606***	0.569***
1 month after treatment	0.358***	0.379***	0.468***	0.424***
2 months after treatment	0.191***	0.213***	0.149*	0.134*
3 months after treatment	0.0584	0.0806**	0.103**	0.0550
4 months after treatment	0.0190	0.0438	0.0678	-0.000425
5 months after treatment	0.1000*	0.124***	0.0743*	0.0428
6 months after treatment	0.0680	0.0906*	0.0664	0.0154
7 months after treatment	0.0201	0.0428	0.0479	-0.00503
8 months after treatment	0.0548	0.0758*	0.0617	0.0299
9 months after treatment	0.0112	0.0328	0.0454	-0.00863
10 months after treatment	0.0531	0.0777*	0.0677	0.0213
11 months after treatment	0.0597	0.0866*	0.0702	0.0508
12 months after treatment	0.0202	0.0469	0.0332	-0.00821
Observations	20,929	20,929	13,539	18,567
R-squared		0.017	0.021	0.020
Number of id	1,001	1,001	500	799

Outpatient Visits

VARIABLES	outpatient_visits Unbalanced No Fixed Effects	outpatient_visits Unbalanced Sample	outpatient_visits Balanced Sample	outpatient_visits IPW Estimation
				*** p<0.01, ** p<0.05, * p<0.1
Fixed Effects	**N**	**Y**	**Y**	**Y**
11 months before treatment	-9.41e-06	-0.0103	0.0109	-0.0179
10 months before treatment	0.0432	0.0233	0.0126	-0.0217
9 months before treatment	0.0308	-0.000481	-0.0174	-0.0442
8 months before treatment	0.0366	-0.00505	0.00149	-0.0685
7 months before treatment	0.0695	0.0156	0.0450	-0.0178
6 months before treatment	0.0645	-0.000470	0.0554	-0.0576
5 months before treatment	0.117*	0.0420	0.151***	0.0161
4 months before treatment	0.259***	0.171***	0.277***	0.127**
3 months before treatment	0.251***	0.151***	0.262***	0.123**
2 months before treatment	0.419***	0.303***	0.364***	0.244***
1 month before treatment	0.953***	0.828***	0.875***	0.786***
Treatment initiation month	3.630***	3.511***	3.008***	3.253***
1 month after treatment	2.478***	2.348***	2.032***	2.269***
2 months after treatment	1.659***	1.520***	1.315***	1.354***
3 months after treatment	1.561***	1.412***	1.426***	1.329***
4 months after treatment	0.995***	0.834***	0.960***	0.776***
5 months after treatment	1.182***	1.010***	0.994***	0.938***
6 months after treatment	1.188***	1.007***	1.116***	0.932***
7 months after treatment	0.739***	0.546***	0.748***	0.506***
8 months after treatment	1.091***	0.888***	0.910***	0.742***
9 months after treatment	1.060***	0.848***	0.950***	0.824***
10 months after treatment	0.765***	0.541***	0.714***	0.517***
11 months after treatment	1.165***	0.931***	0.932***	0.738***
12 months after treatment	0.906***	0.663***	0.866***	0.640***
Observations	20,929	20,929	13,539	18,567
R-squared		0.260	0.287	0.288
Number of id	1,001	1,001	500	799

[10] Although omitted from the output, the regressions on inpatient days and outpatient visits contain the same controls as specified in the body of the paper in the regression on sick days

VARIABLES	sick_days CD4 0-99	sick_days CD4 100-249	sick_days CD4 250-499	sick_days CD4 >=500	sick_days CD4 0-99	sick_days CD4 100-249	sick_days CD4 250-499	sick_days CD4 >=500	sick_days CD4 0-99	sick_days CD4 100-249	sick_days CD4 250-499	sick_days CD4 >=500	sick_days CD4 0-99	sick_days CD4 100-249	sick_days CD4 250-499	sick_days CD4 >=500
Model Type	Unbalanced				Unbalanced				Balanced				IPW			
Fixed Effects	N	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
11 months before treatment	-0.0127	0.0678	-0.056	0.247	0.0268	0.0937	-0.0648	0.252	-0.0745	0.315	-0.077	0.374	-0.0487	0.401	-0.0669	0.584
10 months before treatment	0.46	-0.112	0.035	-0.712	0.524	-0.0649	0.0188	-0.695	0.455	0.116	0.132	-1.295	0.318	-0.0236	0.155	-0.423
9 months before treatment	0.281	-0.0162	0.239	-0.84	0.382	0.063	0.213	-0.842	0.0694	0.354	0.282	-1.351	-0.0754	0.158	0.283	-0.789
8 months before treatment	0.537	0.0964	0.135	-0.316	0.666	0.198	0.099	-0.459	0.313	0.483	-0.0825	-1.359	0.0507	0.328	-0.0459	-0.684
7 months before treatment	-0.231	0.52	0.0655	-0.318	-0.0758	0.650*	0.0193	-0.41	-0.204	0.797**	-0.138	-0.882	-0.083	0.528	0.0067	-0.402
6 months before treatment	0.0907	0.474	-0.145	0.418	0.283	0.634	-0.2	0.32	0.0212	0.891**	-0.337	0.00779	0.215	0.665	-0.189	-0.0952
5 months before treatment	0.265	1.125***	-0.000435	0.245	0.487	1.306***	-0.0626	0.143	0.476	1.414***	-0.00485	-0.211	-0.375	1.307***	0.102	-0.185
4 months before treatment	0.283	1.307***	0.175	0.54	0.543	1.490***	0.1	0.305	0.943	1.792***	0.0887	0.198	-0.00937	1.480***	0.385	0.624
3 months before treatment	1.256	1.179***	0.269	0.0463	1.573**	1.393***	0.189	-0.195	1.309	1.886***	0.482	-0.315	0.715	1.694***	0.906*	0.473
2 months before treatment	2.399***	1.483***	0.469*	0.614	2.739***	1.726***	0.382	0.378	2.978**	2.274***	0.456	0.888	3.150**	2.196***	1.163**	2.579
1 month before treatment	3.040***	2.487***	0.946***	1.196*	3.462***	2.719***	0.849***	1.260*	3.835**	3.422***	0.918**	1.749*	3.510***	3.086***	1.783***	2.799*
Treatment initiation month	10.16***	6.071***	2.820***	2.577***	10.50***	6.387***	2.735***	2.532***	12.00***	6.837***	2.987***	1.897***	10.94***	6.725***	3.783***	2.138**
1 month after treatment	10.55***	4.119***	0.952***	1.689**	10.91***	4.443***	0.856***	1.629***	13.13***	5.065***	0.957**	1.438**	12.77***	5.014***	1.632**	2.631*
2 months after treatment	8.475***	2.001***	0.344	0.376	8.874***	2.361***	0.238	0.318	9.935***	2.785***	0.0985	0.181	9.416***	2.684***	0.2	0.829
3 months after treatment	4.267***	2.025***	0.13	-0.277	4.697***	2.404***	0.016	-0.334	5.974***	2.802***	0.0413	-0.453	4.740***	2.831***	-0.0743	-0.17
4 months after treatment	2.691***	1.782***	0.440*	-0.891	3.132***	2.181***	0.322	-0.949	4.287***	2.448***	0.418	-1.043	2.740*	2.463***	0.226	-0.354
5 months after treatment	2.256**	1.519***	0.369	-0.602	2.713***	1.949***	0.24	-0.666	3.611**	2.154***	0.0318	-0.778	1.717	2.197***	-0.0219	-0.145
6 months after treatment	0.964	1.286***	0.627**	-0.877	1.428*	1.743**	0.486	-0.942	1.571	1.991***	0.0907	-0.852	0.16	1.987**	0.197	-0.353
7 months after treatment	0.44	1.188***	0.00774	-0.0232	0.978	1.669**	-0.134	-0.0787	0.846	2.364***	-0.558	-0.282	-0.443	1.948**	-0.377	0.282
8 months after treatment	0.763	1.518***	0.0382	-0.395	1.320*	2.026***	-0.114	-0.434	0.945	2.705***	-0.385	-0.321	0.411	2.402***	-0.233	-0.0352
9 months after treatment	0.703	1.062**	-0.032	0.33	1.284	1.593**	-0.189	0.274	1.357	2.176***	-0.343	-0.56	0.609	1.971**	-0.334	0.418
10 months after treatment	0.352	0.732*	-0.132	0.397	0.948	1.267*	-0.294	0.321	0.138	1.679***	-0.811	-0.671	0.303	1.458*	-0.476	0.426
11 months after treatment	-0.385	0.768*	0.287	-0.214	0.224	1.327*	0.116	-0.345	0.537	1.763***	-0.19	-0.532	-0.809	1.543**	-0.195	0.342
12 months after treatment	-0.609	0.805*	0.188	0.183	0.0141	1.393*	0.00742	-0.00656	0.161	1.887***	-0.295	-0.618	-1.255	1.749**	-0.274	0.118
Observations	2,382	6,921	9,114	1,333	2,382	6,921	9,114	1,333	1,468	4,726	6,003	1,044	2,104	6,669	8,133	1,065
R-squared					0.21	0.085	0.038	0.053	0.26	0.096	0.054	0.077	0.284	0.096	0.074	0.116

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Summary Statistics

Year	Negative	Positive	Positive as a Percentage of Tested	Unknown	Total
2009	7,395	1,248	1,248	667	9,310
	79.43%	13.40%	14.43%	7.16%	100.00%
2010	7,636	1,331	1,331	687	9,654
	79.10%	13.79%	14.85%	7.12%	100.00%
2011	7,596	1,350	1,350	478	9,424
	80.60%	14.33%	15.10%	5.07%	100.00%
2012	7,420	1,373	1,373	411	9,204
	80.62%	14.92%	15.62%	4.47%	100.00%
2013	7,521	1,400	1,400	291	9,212
	81.64%	15.20%	15.70%	3.16%	100.00%
2014	7,184	1,380	1,380	211	8,775
	81.87%	15.73%	16.12%	2.40%	116.12%
2015	6,958	1,337	1,337	141	8,436
	82%	16%	16%	2%	100%
2016	5,899	1,168	1,168	89	7,156
	82.43%	16.32%	16.53%	1.24%	100.00%

year	Untreated	First Round	Second Round	Third Round	Fourth Round	Total Treated	Total
2009	343	1065	61	0	0	1126	1,469
	23.35%	72.50%	4.15%	0.00%	0.00%	76.65%	100.00%
2010	371	1057	79	8	0	1144	1,515
	24.49%	69.77%	5.21%	0.53%	0.00%	75.51%	100.00%
2011	355	1016	92	16	0	1124	1,479
	24.00%	68.70%	6.22%	1.08%	0.00%	76.00%	100.00%
2012	372	974	101	23	0	1098	1,470
	25.31%	66.26%	6.87%	1.56%	0.00%	74.69%	100.00%
2013	395	922	123	24	7	1076	1,471
	26.85%	62.68%	8.36%	1.63%	0.48%	73.15%	100.00%
2014	410	836	140	22	8	1006	1,416
	28.95%	59.04%	9.89%	1.55%	0.56%	71.04%	100.00%
2015	410	765	142	30	6	943	1,353
	30%	57%	11%	2%	0%	69.70%	100%
2016	380	645	113	28	4	790	1,170
	32.48%	55.13%	9.66%	2.39%	0.34%	67.52%	100.00%

Bibliography

Asfaw, A., Ali, D., Eticha, T., Alemayehu, A., Alemayehu, M., & Kindeya, F. (2015). CD4 cell count trends after commencement of antiretroviral therapy among HIV-infected patients in Tigray, Northern Ethiopia: A retrospective cross-sectional study. *PLoS ONE*, *10*(3), 1–9. https://doi.org/10.1371/journal.pone.0122583

Attia, S., Egger, M., Müller, M., Zwahlen, M., & Low, N. (2009). Sexual transmission of HIV according to viral load and antiretroviral therapy: systematic review and meta-analysis. *AIDS (London, England)*, *23*(11), 1397–404. https://doi.org/10.1097/QAD.0b013e32832b7dca

Bärnighausen, T., Bor, J., Wandira-Kazibwe, S., & Canning, D. (2011). Correcting HIV prevalence estimates for survey nonparticipation using heckman-type selection models. *Epidemiology*, *22*(1), 27–35. https://doi.org/10.1097/EDE.0b013e3181ffa201

Battegay, M., Nüesch, R., Hirschel, B., & Kaufmann, G. R. (2006). Immunological recovery and antiretroviral therapy in HIV-1 infection. *Lancet Infectious Diseases*, *6*(5), 280–287. https://doi.org/10.1016/S1473-3099(06)70463-7

Bor, J., Tanser, F., Newell, M. L., & Bärnighausen, T. (2012). In a study of a population cohort in South Africa, HIV patients on antiretrovirals had nearly full recovery of employment. *Health Affairs*, *31*(7), 1459–1469. https://doi.org/10.1377/hlthaff.2012.0407

Charalambous, S., Innes, C., Muirhead, D., Kumaranayake, L., Fielding, K., Pemba, L., … Churchyard, G. J. (2007). Evaluation of a workplace HIV treatment programme in South Africa. *Aids*, *21*(SUPPL. 3), S73–S78. https://doi.org/10.1097/01.aids.0000279696.63438.aa

Charalambous, S., Innes, C., Muirhead, D., Kumaranayake, L., Fielding, K., Pemba, L., … Churchyard, G. J. (2007). Evaluation of a workplace HIV treatment programme in South Africa. *Aids*, *21*(SUPPL. 3), 73–78. https://doi.org/10.1097/01.aids.0000279696.63438.aa

Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C., & Dye, C. (2003). The growing burden of tuberculosis: Global trends and interactions with the HIV epidemic. *Archives of Internal Medicine*. https://doi.org/10.1001/archinte.163.9.1009

Deeks, S. G., Lewin, S. R., & Havlir, D. V. (2013). The end of AIDS: HIV infection as a chronic disease. *The Lancet*. https://doi.org/10.1016/S0140-6736(13)61809-7

Dunn, D. (2006). Predictive value of absolute CD4 cell count for disease progression in untreated HIV-1-infected children. *AIDS*, *20*(9), 1289–1294. https://doi.org/10.1097/01.aids.0000232237.20792.68

Fox, M. P., Rosen, S., MacLeod, W. B., Wasunna, M., Bii, M., Foglia, G., & Simon, J. L. (2004). The impact of HIV/AIDS on labour productivity in Kenya. *Tropical Medicine and International Health*, *9*(3), 318–324. https://doi.org/10.1111/j.1365-3156.2004.01207.x

Freedberg, K. A., Losina, E., Weinstein, M. C., Paltiel, A. D., Cohen, C. J., Seage, G. R., … Goldie, S. J. (2001). The cost effectiveness of combination antiretroviral therapy for HIV disease. *The New England Journal of Medicine*, *344*(11), 824–831. https://doi.org/10.1056/NEJM200103153441108

Goldstein, B., & Webster, I. (1972). COAL WORKERS' PNEUMOCONIOSIS IN SOUTH AFRICA. *Annals of the New York Academy of Sciences*, *200*(1), 306–315. https://doi.org/10.1111/j.1749-6632.1972.tb40197.x

Granich, R., Kahn, J. G., Bennett, R., Holmes, C. B., Garg, N., Serenata, C., … Williams, B. G. (2012). Expanding ART for treatment and prevention of HIV in South Africa: Estimated cost and cost-effectiveness 2011-2050. *PLoS ONE*, *7*(2). https://doi.org/10.1371/journal.pone.0030216

Habyarimana, J., Mbakile, B., & Pop-Eleches, C. (2010). The Impact of HIV/AIDS and ARV Treatment on Worker Absenteeism: Implications for African Firms. *Journal of Human Resources*, *45*(February 2007), 809–839. https://doi.org/10.1353/jhr.2010.0032

Hogg, R. S. (2001). Rates of Disease Progression by Baseline CD4 Cell Count and Viral Load After Initiating Triple-Drug Therapy. *JAMA*, *286*(20), 2568. https://doi.org/10.1001/jama.286.20.2568

Holmes, C. B., Wood, R., Badri, M., Zilber, S., Wang, B., Maartens, G., … Losina, E. (2006). CD4 decline and incidence of opportunistic infections in Cape Town, South Africa: implications for prophylaxis and treatment. *J Acquir Immune Defic Syndr*, *42*(4), 464–469. https://doi.org/10.1097/01.qai.0000225729.79610.b7

INSIGHT START Study Group. (2015). Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. *The New England Journal of Medicine*, *373*(9), 795–807. https://doi.org/10.1056/NEJMoa1506816

Jaffar, S., Wilkins, A., Ngom, P. T., Sabally, S., Corrah, T., Bangali, J. E., … Whittle, H. C. (1997). Rate of decline of percentage CD4+ cells is faster in HIV-1 than in HIV-2 infection. *J Acquir Immune Defic Syndr Hum Retrovirol*, *16*(5), 327–332. https://doi.org/10.1097/00042560-199712150-00003

Kim, Y. M., Kim, M., Kim, S. K., Park, K., Jin, S. H., Lee, U. S., … Lee, S. B. (2009). Mycobacterial infections in coal workers' pneumoconiosis patients in South Korea. *Scandinavian Journal of Infectious Diseases*, *41*(9), 656–662. https://doi.org/10.1080/00365540903089468

Larson, B. A., Fox, M. P., Bii, M., Rosen, S., Rohr, J., Shaffer, D., … Simon, J. L. (2013). Antiretroviral therapy, labor productivity, and sex: A longitudinal cohort study of tea pluckers in Kenya. *Aids*, *27*(1), 115–123. https://doi.org/10.1097/QAD.0b013e32835a5b12

Maduna, P. H., Dolan, M., Kondlo, L., Mabuza, H., Dlamini, J. N., Polis, M., … Neaton, J. D. (2015). Morbidity and mortality according to latest CD4+ cell count among HIV positive individuals in South Africa who enrolled in Project Phidisa. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0121843

Menzies, N. A., Berruti, A. A., Berzon, R., Filler, S., Ferris, R., Ellerbrock, T. V., & Blandford, J. M. (2011). The cost of providing comprehensive HIV treatment in PEPFAR-supported programs. *AIDS*, *25*(14), 1753–1760. https://doi.org/10.1097/QAD.0b013e3283463eec

Meyer-Rath, G., Pienaar, J., Brink, B., van Zyl, A., Muirhead, D., Grant, A.,

… Vickerman, P. (2015). The Impact of Company-Level ART Provision to a Mining Workforce in South Africa: A Cost–Benefit Analysis. *PLoS Medicine*, *12*(9). https://doi.org/10.1371/journal.pmed.1001869

Meyer-Rath, G., Pienaar, J., Brink, B., Van Zyl, A., Muirhead, D., Grant, A., … Vickerman, P. (2012). Company-level ART provision to employees is cost saving: A modelled cost-benefit analysis of the impact of HIV and antiretroviral treatment in a mining workforce in South Africa. *Journal of the International AIDS Society*, *15*, 274.

Mocroft, A., Phillips, A., Gatell, J., Ledergerber, B., Fisher, M., Clumeck, N., … Lundgren, J. (2007). Normalisation of CD4 counts in patients with HIV-1 infection and maximum virological suppression who are taking combination antiretroviral therapy: an observational cohort study. *Lancet*, *370*(9585), 407–413. https://doi.org/10.1016/S0140-6736(07)60948-9

Morgan, D., Mahe, C., Mayanja, B., Okongo, J. M., Lubega, R., & Whitworth, J. A. G. (2002). HIV-1 infection in rural Africa: Is there a difference in median time to AIDS and survival compared with that in industrialized countries? *Aids*, *16*(4), 597–603. https://doi.org/10.1097/00002030-200203080-00011

Muirhead, D., Kumaranayake, L., Hongoro, C., Charalambous, S., Grant, A., Fielding, K., & Churchyard, G. (2006). DFID SRH & HIV RPC at the XVI International AIDS Conference, Toronto, 2006, 2006.

Okulicz, J. F., Le, T. D., Agan, B. K., Camargo, J. F., Landrum, M. L., Wright, E., … Ahuja, S. K. (2015). Influence of the timing of antiretroviral therapy on the potential for normalization of immune status in human immunodeficiency virus 1-infected individuals. *JAMA Internal Medicine*, *175*(1), 88–99. https://doi.org/10.1001/jamainternmed.2014.4010

Robles, R. R., Colón, H. M., Díaz, N., Cancel, L. I., MacGowan, R., Cole, G. E., & Allen, D. M. (1994). Behavioural risk factors and HIV infection of injection drug users at detoxification clinics in Puerto Rico. *International Journal of Epidemiology*, *23*(3), 595–601.

Sharma, S. K., & Soneja, M. (2011). HIV & immune reconstitution inflammatory syndrome (IRIS). *The Indian Journal of Medical Research*, *134*(December), 866–77. https://doi.org/10.4103/0971-5916.92632

Shisana, O., Rehle, T., Simbayi, L. C., Zuma, K., Jooste, S., Zungu, N., … Al., E. (2014). South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. *HSRC Press*, 194. https://doi.org/10.4314/ajpsy.v13i4.61877

Sonnenberg, P., Copas, A., Glynn, J. R., Bester, A., Nelson, G., Shearer, S., & Murray, J. (2011). The effect of HIV infection on time off work in a large cohort of gold miners with known dates of seroconversion. *Occupational and Environmental Medicine*, *68*(9), 647–652. https://doi.org/10.1136/oem.2010.058487

Weuve, J., Tchetgen Tchetgen, E. J., Glymour, M. M., Beck, T. L., Aggarwal, N. T., Wilson, R. S., … Mendes De Leon, C. F. (2012). Accounting for bias due to selective attrition: The example of smoking and cognitive decline. *Epidemiology*, *23*(1), 119–128. https://doi.org/10.1097/EDE.0b013e318230e861

Williams, B. G., Hargrove, J. W., & Humphrey, J. H. (2010). The benefits of early treatment for HIV. *AIDS*. https://doi.org/10.1097/QAD.0b013e32833ac860

Wooldridge, J. M. (2002). Inverse probability weighted M-estimators for sample selection, attrition, and stratification. *Portuguese Economic Journal*, *1*(2), 117–139. https://doi.org/10.1007/s10258-002-0008-x

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics*, *141*(2), 1281–1301. https://doi.org/10.1016/j.jeconom.2007.02.002